

WHAT IS CLAIMED IS:

1. A system, comprising:

a psycho-physical state detection mechanism for detecting psycho-physical state of a user based on the speech from the user; and

5 a spoken dialogue mechanism for carrying on a dialogue with said user based on the psycho-physical state of the user, detected by the psycho-physical detection mechanism from the speech from the user.

2. The system according to claim 1, wherein said spoken dialogue mechanism comprises:

a speech understanding mechanism for understanding the speech from the user based on the psycho-physical state of the user to generate a literal meaning of the speech; and

a voice response generation mechanism for generating a voice response to the user based on the literal meaning of the speech and the psycho-physical state of the user.

3. The system according to claim 2, wherein said speech understanding mechanism comprises:

at least one acoustic model for characterizing the acoustic properties of speech, each of said at least one acoustic model corresponding to some distinct characteristic related to a psycho-physical state of a speaker;

an acoustic model selection mechanism for selecting an acoustic model that is appropriate to according to the psycho-physical state detected by the psycho-physical state detection mechanism;

a speech recognizer for generating a transcription of spoken words recognized from the speech using the acoustic model selected by the acoustic model selection mechanism; and  
a language understanding mechanism for interpreting the literal meaning of the speech based on the transcription.

5

4. The system according to claim 2, wherein said voice response generation mechanism comprises:

a natural language response generator for generating a response based on an understanding of the transcription, said response being generated appropriately according to the psycho-physical state of the user;

a prosodic pattern determining mechanism for determining the prosodic pattern to be applied to said response that is considered as appropriate according to the psycho-physical state; and

a text-to-speech engine for synthesizing the voice response based on said response and said prosodic pattern.

5. The system according to claim 1, wherein said psycho-physical state detection mechanism comprises:

an acoustic feature extractor for extracting acoustic features from input speech data to generate at least one acoustic feature; and

a psycho-physical state classifier for classifying the input speech data into one or more psycho-physical states based on said at least one acoustic feature.

6. The system according to claim 5, further comprising:

at least one psycho-physical state model, each of said at least one psycho-physical state model corresponding to a single psycho-physical state and characterizing the acoustic properties of the single psycho-physical state; and

5 an off-line training mechanism for establishing said at least one psycho-physical model based on labeled training speech data.

7. The system according to claim 1, further comprising a dialogue manager that control the dialogue flow.

8. A voice based information retrieval system, comprising:

an information database for archive information, said information being accessible and retrievable;

a search engine for accessing and retrieving said information stored in the information database; and

a psycho-physical state sensitive spoken dialogue system connecting to the search engine and a user, voice communicating with the user in a psycho-physical state sensitive manner, responding to the user's request for desired information by activating the search engine to retrieve the desired information, and generating a voice response to the user according to the desired information and the detected psycho-physical state of the user.

9. The system according to claim 8, wherein said information database includes at least one domain information database, each of the domain information database storing the information related to at least one specific domain of interest.

5 10. A method, comprising:  
receiving, by a psycho-physical state detection mechanism, input speech data from a user;  
detecting the psycho-physical state of the user from the input speech data;  
understanding, by a speech understanding mechanism, the literal meaning of spoken  
10 words recognized from the input speech data based on the psycho-physical state of the user,  
detected by said detecting; and  
generating, by a voice response generation mechanism, a voice response to the user  
based on the literal meaning of the input speech data and the psycho-physical state of the user.

15 11. The method according to claim 10, wherein said detecting comprises:  
extracting, by a acoustic feature extractor, at least one acoustic feature from the input  
speech data; and  
classifying, by a psycho-physical state classifier and based on said at least one feature,  
the input speech data into the psycho-physical state according to at least one psycho-physical  
20 state model.

12. The method according to claim 11, further comprising:

receiving, by an off-line training mechanism, labeled training data, wherein each of the data items in said labeled training data is labeled by a psycho-physical state; and

building said at least one psycho-physical state model using the labeled training data, each of the at least one psycho-physical state model corresponding to a single psycho-physical state and being established based on the data items in the labeled training data that have a label corresponding to the single psycho-physical state.

13. The method according to claim 10, wherein said understanding comprises:

selecting, by an acoustic model selection mechanism, an acoustic model, from at least one acoustic model, that is appropriate to according to the psycho-physical state, detected by said detecting, each of said at least one acoustic model corresponding to some distinct speech characteristic related to a psycho-physical state;

recognizing, by a speech recognizer, the spoken words from the input speech data using the acoustic model, selected by said selecting, to generate a transcription; and

interpreting, by a language understanding mechanism, the literal meaning of the spoken words based on the transcription.

14. The method according to claim 10, wherein said generating comprises:

constructing, by a natural language response generator, a natural language response based on an understanding of the transcription, said natural language response being constructed appropriately according to the psycho-physical state of the user;

determining, by a prosodic pattern determining mechanism, the prosodic pattern to be applied to said natural language response, wherein the prosodic pattern is considered to be appropriate according to the psycho-physical state; and

synthesizing, by a text-to-speech engine, the voice response based on said natural

5 language response and said prosodic pattern.

15. A method for voice based information retrieval, comprising:

communicating between a psycho-physical state sensitive spoken dialogue system and a user via voice to understand the user's request for desired information, wherein said  
10 understand is achieved according to the psycho-physical state of the user;

retrieving, by a search engine, information from an information database based on the understanding of the user's request for desired information to generate retrieved information;  
and

generating, by the psycho-physical state sensitive spoken dialogue system, a voice  
15 response to the user's request based on the retrieved information and the psycho-physical state of the user.

16. The method according to claim 15, wherein said communicating comprises:

receiving input speech data from the user;

20 detecting the psycho-physical state of the user from the input speech data; and

recognizing the user's request based on the psycho-physical state of the user.

17. The method according to claim 15, wherein said desired information includes information about at least one of:

weather;  
restaurants,  
news;  
sports;  
movies;  
stocks; and  
driving directions.

18. A computer-readable medium encoded with a program, said program comprising: receiving, by a psycho-physical state detection mechanism, input speech data from a user;

detecting the psycho-physical state of the user from the input speech data;

understanding, by a speech understanding mechanism, the literal meaning of spoken words recognized from the input speech data based on the psycho-physical state of the user, detected by said detecting; and

generating, by a voice response generation mechanism, a voice response to the user based on the literal meaning of the input speech data and the psycho-physical state of the user.

19. The medium according to claim 18, wherein said detecting comprises:

extracting, by a acoustic feature extractor, at least one acoustic feature from the input speech data; and

classifying, by a psycho-physical state classifier and based on said at least one feature, the input speech data into the psycho-physical state according to at least one psycho-physical state model.

5           20. The medium according to claim 19, further comprising:  
  
receiving, by an off-line training mechanism, labeled training data, wherein each of the data items in said labeled training data is labeled by a psycho-physical state; and  
  
building said at least one psycho-physical state model using the labeled training data, each of the at least one psycho-physical state model corresponding to a single psycho-physical  
10 state and being established based on the data items in the labeled training data that have a label corresponding to the single psycho-physical state.

21. The medium according to claim 18, wherein said understanding comprises:  
  
selecting, by an acoustic model selection mechanism, an acoustic model, from at least  
15 one acoustic model, that is appropriate to according to the psycho-physical state, detected by said detecting, each of said at least one acoustic model corresponding to some distinct speech characteristic related to a psycho-physical state;

recognizing, by a speech recognizer, the spoken words from the input speech data using the acoustic model, selected by said selecting, to generate a transcription; and

20           interpreting, by a language understanding mechanism, the literal meaning of the spoken words based on the transcription.

22. The medium according to claim 18, wherein said generating comprises:



constructing, by a natural language response generator, a natural language response based on an understanding of the transcription, said natural language response being constructed appropriately according to the psycho-physical state of the user;

determining, by a prosodic pattern determining mechanism, the prosodic pattern to be applied to said natural language response, wherein the prosodic pattern is considered to be appropriate according to the psycho-physical state; and

synthesizing, by a text-to-speech engine, the voice response based on said natural language response and said prosodic pattern.

23. A computer-readable medium encoded with a program for voice based information retrieval, said program comprising:

communicating between a psycho-physical state sensitive spoken dialogue system and a user via voice to understand the user's request for desired information, wherein said understand is achieved according to the psycho-physical state of the user;

retrieving, by a search engine, information from an information database based on the understanding of the user's request for desired information to generate retrieved information; and

generating, by the psycho-physical state sensitive spoken dialogue system, a voice response to the user's request based on the retrieved information and the psycho-physical state of the user.

24. The medium according to claim 23, wherein said communicating comprises: receiving input speech data from the user;

Parameter	1990-1991		1991-1992		1992-1993		1993-1994		1994-1995		1995-1996		1996-1997		1997-1998		1998-1999		1999-2000		2000-2001		2001-2002		2002-2003		2003-2004		2004-2005		2005-2006		2006-2007		2007-2008		2008-2009		2009-2010		2010-2011		2011-2012		2012-2013		2013-2014		2014-2015		2015-2016		2016-2017		2017-2018		2018-2019		2019-2020		2020-2021		2021-2022		2022-2023		2023-2024		2024-2025		2025-2026		2026-2027		2027-2028		2028-2029		2029-2030		2030-2031		2031-2032		2032-2033		2033-2034		2034-2035		2035-2036		2036-2037		2037-2038		2038-2039		2039-2040		2040-2041		2041-2042		2042-2043		2043-2044		2044-2045		2045-2046		2046-2047		2047-2048		2048-2049		2049-2050		2050-2051		2051-2052		2052-2053		2053-2054		2054-2055		2055-2056		2056-2057		2057-2058		2058-2059		2059-2060		2060-2061		2061-2062		2062-2063		2063-2064		2064-2065		2065-2066		2066-2067		2067-2068		2068-2069		2069-2070		2070-2071		2071-2072		2072-2073		2073-2074		2074-2075		2075-2076		2076-2077		2077-2078		2078-2079		2079-2080		2080-2081		2081-2082		2082-2083		2083-2084		2084-2085		2085-2086		2086-2087		2087-2088		2088-2089		2089-2090		2090-2091		2091-2092		2092-2093		2093-2094		2094-2095		2095-2096		2096-2097		2097-2098		2098-2099		2099-2100		2100-2101		2101-2102		2102-2103		2103-2104		2104-2105		2105-2106		2106-2107		2107-2108		2108-2109		2109-2110		2110-2111		2111-2112		2112-2113		2113-2114		2114-2115		2115-2116		2116-2117		2117-2118		2118-2119		2119-2120		2120-2121		2121-2122		2122-2123		2123-2124		2124-2125		2125-2126		2126-2127		2127-2128		2128-2129		2129-2130		2130-2131		2131-2132		2132-2133		2133-2134		2134-2135		2135-2136		2136-2137		2137-2138		2138-2139		2139-2140		2140-2141		2141-2142		2142-2143		2143-2144		2144-2145		2145-2146		2146-2147		2147-2148		2148-2149		2149-2150		2150-2151		2151-2152		2152-2153		2153-2154		2154-2155		2155-2156		2156-2157		2157-2158		2158-2159		2159-2160		2160-2161		2161-2162		2162-2163		2163-2164		2164-2165		2165-2166		2166-2167		2167-2168		2168-2169		2169-2170		2170-2171		2171-2172		2172-2173		2173-2174		2174-2175		2175-2176		2176-2177		2177-2178		2178-2179		2179-2180		2180-2181		2181-2182		2182-2183		2183-2184		2184-2185		2185-2186		2186-2187		2187-2188		2188-2189		2189-2190		2190-2191		2191-2192		2192-2193		2193-2194		2194-2195		2195-2196		2196-2197		2197-2198		2198-2199		2199-2200		2200-2201		2201-2202		2202-2203		2203-2204		2204-2205		2205-2206		2206-2207		2207-2208		2208-2209		2209-2210		2210-2211		2211-2212		2212-2213		2213-2214		2214-2215		2215-2216		2	
-----------	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	---	--

1. The first part of the document is a title page. It contains the title of the document, the author's name, and the date of the document.